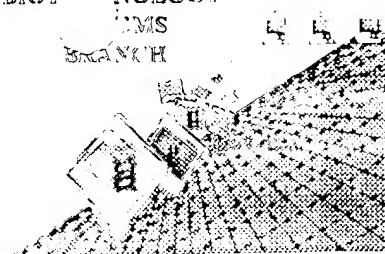# RAW SEQUENCE LISTING ERROR REPORT

The Biotechnology Systems Branch of the Scientific and Technical Information Center (STIC) detected errors when processing the following computer readable form:

Application Serial Number: 09/719,379

Source: PCT09

Date Processed by STIC: 2-15-2001

THE ATTACHED PRINTOUT EXPLAINS DETECTED ERRORS.
PLEASE FORWARD THIS INFORMATION TO THE APPLICANT BY EITHER:
1) INCLUDING A COPY OF THIS PRINTOUT IN YOUR NEXT COMMUNICATION TO THE APPLICANT, WITH A NOTICE TO COMPLY or,
2) TELEPHONING APPLICANT AND FAXING A COPY OF THIS PRINTOUT, WITH A NOTICE TO COMPLY
FOR CRF SUBMISSION QUESTIONS, PLEASE CONTACT MARK SPENCER, 703-308-4212.

FOR SEQUENCE RULES INTERPRETATION, PLEASE CONTACT ROBERT WAX, 703-308-4216.
PATENTIN 2.1 e-mail help: patin21help@uspto.gov or phone 703-306-4119 (R. Wax)
PATENTIN 3.0 e-mail help: patin3help@uspto.gov or phone 703-306-4119 (R. Wax)

TO REDUCE ERRORED SEQUENCE LISTINGS, PLEASE USE THE CHECKER VERSION 3.0 PROGRAM, ACCESSIBLE THROUGH THE U.S. PATENT AND TRADEMARK OFFICE WEBSITE. SEE BELOW:

Checker Version 3.0

The Checker Version 3.0 application is a state-of the-art Windows based software program employing a logical and intuitive user-interface to check whether a sequence listing is in compliance with format and content rules. Checker Version 3.0 works for sequence listings generated for the original version of 37 CFR §§1.821 – 1.825 effective October 1, 1990 (old rules) and the revised version (new rules) effective July 1, 1998 as well as World Intellectual Property Organization (WIPO) Standard ST.25.
Checker Version 3.0 replaces the previous DOS-based version of Checker, and is Y2K-compliant. Checker allows public users to check sequence listings in Computer Readable form (CRF) before submitting them to the United States Patent and Trademark Office (USPTO).
Use of Checker prior to filing the sequence listing is expected to result in fewer errored sequence listings, thus saving time and money.

Checker Version 3.0 can be down loaded from the USPTO website at the following address:
http://www.uspto.gov/web/offices/pac/checker

R●Sequence Listing Error S●mary

ERROR DETECTED | SUGGESTED CORRECTION

SERIAL NUMBER: 09/719,379

ATTN: NEW RULES CASES: PLEASE DISREGARD ENGLISH "ALPHA" HEADERS, WHICH WERE INSERTED BY PTO SOFTWARE

1 _____ Wrapped Nucleics — The number/text at the end of each line "wrapped" down to the next line. This may occur if your file was retrieved in a word processor after creating it. Please adjust your right margin to .3, as this will prevent "wrapping".

2 _____ Wrapped Aminos — The amino acid number/text at the end of each line "wrapped " down to the next line. This may occur if your file was retrieved in a word processor after creating it. Please adjust your right margin to .3, as this will prevent "wrapping".

3 _____ Incorrect Line Length — The rules require that a line not exceed 72 characters in length. This includes spaces.

4 _____ Misaligned Amino Acid Numbering — The numbering under each 5th amino acid is misaligned. This may be caused by the use of tabs between the numbering. It is recommended to delete any tabs and use spacing between the numbers.

5 _____ Non-ASCII — This file was not saved in ASCII (DOS) text, as required by the Sequence Rules. Please ensure your subsequent submission is saved in ASCII text so that it can be processed.

6 _____ Variable Length — Sequence(s) _____ contain n's or Xaa's which represented more than one residue. As per the rules, each n or Xaa can only represent a single residue. Please present the maximum number of each residue having variable length and indicate in the (ix) feature section that some may be missing.

7 _____ PatentIn ver. 2.0 "bug" — A "bug" in PatentIn version 2.0 has caused the <220>-<223> section to be missing from amino acid sequence(s) _____. Normally, PatentIn would automatically generate this section from the previously coded nucleic acid sequence. Please manually copy the relevant <220>-<223> section to the subsequent amino acid sequence.

8 _____ Skipped Sequences (OLD RULES) — Sequence(s) _____ missing. If intentional, please use the following format for each skipped sequence:
(2) INFORMATION FOR SEQ ID NO:X:
(i) SEQUENCE CHARACTERISTICS:(Do not insert any headings under "SEQUENCE CHARACTERISTICS")
(xi) SEQUENCE DESCRIPTION:SEQ ID NO:X:
This sequence is intentionally skipped

Please also adjust the "(iii) NUMBER OF SEQUENCES:" response to include the skipped sequence(s).

9 _____ Skipped Sequences (NEW RULES) — Sequence(s) _____ missing. If intentional, please use the following format for each skipped sequence.
<210> sequence Id number
<400> sequence Id number
000

10 ✓ Use of n's or Xaa's (NEW RULES) — Use of n's and/or Xaa's have been detected in the Sequence Listing. Use of <220> to <223> is MANDATORY if n's or Xaa's are present. In <220> to <223> section, please explain location of n or Xaa, and which residue n or Xaa represents.

11 _____ Use of <213>Organism (NEW RULES) — Sequence(s) _____ are missing this mandatory field or its response.

12 ✓ Use of <220>Feature (NEW RULES) — Sequence(s) 73 are missing the <220>Feature and associated headings. Use of <220> to <223> is MANDATORY if <213>ORGANISM is "Artificial" or "Unknown". Please explain source of genetic material in <220> to <223> section.
(See "Federal Register," 6/01/98, Vol. 63, No. 104, pp. 29631-32) (Sec. 1.823 of new Rules)

13 _____ PatentIn ver. 2.0 "bug" — Please do not use "Copy to Disk" function of PatentIn version 2.0. This causes a corrupted file, resulting in missing mandatory numeric identifiers and responses (as indicated on raw sequence listing). Instead, please use "File Manager" or any other means to copy file to floppy disk.

AKS-Biotechnology Systems Branch- 5/15/99

PCT09

```
            RAW SEQUENCE LISTING                DATE: 02/15/2001
            PATENT APPLICATION:  US/09/719,379  TIME: 14:29:54

            Input Set : A:\seqlist.txt
            Output Set: N:\CRF3\02152001\I719379.raw

    4 <110> APPLICANT: Joe Cohen
    5          Yves Lobet
    6          Guy Dequesne
    7          Lauren Bakaletz
    9 <120> TITLE OF INVENTION: Vaccine
   12 <130> FILE REFERENCE: B45145
C--> 14 <140> CURRENT APPLICATION NUMBER: US/09/719,379
C--> 14 <141> CURRENT FILING DATE: 2000-12-11
   14 <160> NUMBER OF SEQ ID NOS: 81
   16 <170> SOFTWARE: FastSEQ for Windows Version 3.0
   18 <210> SEQ ID NO: 1
   19 <211> LENGTH: 19
   20 <212> TYPE: PRT
   21 <213> ORGANISM: Haemophilus influenzae strain ntHi-10567RM
   23 <400> SEQUENCE: 1
   24  Arg Ser Asp Tyr Lys Phe Tyr Glu Ala Ala Asn Gly Thr Arg Asp His
   25  1               5                  10                  15
   26  Lys Lys Gly
   29 <210> SEQ ID NO: 2
   30 <211> LENGTH: 22
   31 <212> TYPE: PRT
   32 <213> ORGANISM: Haemophilus influenzae strain ntHi-1715MEE
   34 <400> SEQUENCE: 2
   35  Arg Ser Asp Tyr Lys Leu Tyr Asn Lys Asn Ser Ser Ser Asn Ser Thr
   36  1               5                  10                  15
   37  Leu Lys Asn Leu Gly Glu
   38             20
   40 <210> SEQ ID NO: 3
   41 <211> LENGTH: 13
   42 <212> TYPE: PRT
   43 <213> ORGANISM: Haemophilus influenzae strain ntHi-1729MEE
   45 <400> SEQUENCE: 3
   46  Arg Ser Asp Tyr Lys Phe Tyr Asp Asn Lys Arg Ile Asp
   47  1               5                  10
   49 <210> SEQ ID NO: 4
   50 <211> LENGTH: 19
   51 <212> TYPE: PRT
   52 <213> ORGANISM: Haemophilus influenzae strain ntHi-183NP
   54 <400> SEQUENCE: 4
   55  Arg Ser Asp Tyr Lys Leu Tyr Asn Lys Asn Ser Ser Thr Leu Lys Asp
   56  1               5                  10                  15
   57  Leu Gly Glu
   60 <210> SEQ ID NO: 5
   61 <211> LENGTH: 19
   62 <212> TYPE: PRT
   63 <213> ORGANISM: Haemophilus influenzae strain ntHi-1128
   65 <400> SEQUENCE: 5
```

Does Not Comply
Corrected Diskette Needed

```
 66   Arg Ser Asp Tyr Lys Phe Tyr Glu Asp Ala Asn Gly Thr Arg Asp His
 67   1               5                   10                  15
 68   Lys Lys Gly
 71   <210> SEQ ID NO: 6
 72   <211> LENGTH: 19
 73   <212> TYPE: PRT
 74   <213> ORGANISM: Non-typable Haemophilus influenzae
 76   <400> SEQUENCE: 6
```

W-->
```
 77   Arg Ser Asp Tyr Lys Phe Tyr Glu Ala Pro Asn Ser Thr Arg Asp Xaa
 78   1               5                   10                  15
 79   Lys Lys Gly
 82   <210> SEQ ID NO: 7
 83   <211> LENGTH: 19
 84   <212> TYPE: PRT
 85   <213> ORGANISM: Haemophilus influenzae strain ntHi-152NP
 87   <400> SEQUENCE: 7
 88   Arg Ser Asp Tyr Lys Phe Tyr Glu Asp Ala Asp Gly Thr Arg Asp His
 89   1               5                   10                  15
 90   Lys Lys Gly
 93   <210> SEQ ID NO: 8
 94   <211> LENGTH: 19
 95   <212> TYPE: PRT
 96   <213> ORGANISM: Haemophilus influenzae strain ntHi-1234MEE
 98   <400> SEQUENCE: 8
 99   Arg Ser Asp Tyr Lys Phe Tyr Asp Asp Ala Asn Gly Thr Arg Asp His
100   1               5                   10                  15
101   Lys Lys Gly
104   <210> SEQ ID NO: 9
105   <211> LENGTH: 19
106   <212> TYPE: PRT
107   <213> ORGANISM: Haemophilus influenzae strain ntHi-90100RM
109   <400> SEQUENCE: 9
110   Arg Ser Asp Tyr Lys Phe Tyr Glu Asp Glu Asn Gly Thr Arg Asp His
111   1               5                   10                  15
112   Lys Lys Gly
115   <210> SEQ ID NO: 10
116   <211> LENGTH: 19
117   <212> TYPE: PRT
118   <213> ORGANISM: Haemophilus influenzae strain ntHi-86027NP
120   <400> SEQUENCE: 10
121   Arg Ser Asp Tyr Lys Phe Tyr Glu Val Ala Asn Gly Thr Arg Asp His
122   1               5                   10                  15
123   Lys Lys Gly
126   <210> SEQ ID NO: 11
127   <211> LENGTH: 19
128   <212> TYPE: PRT
129   <213> ORGANISM: Haemophilus influenzae strain ntHi-476
131   <400> SEQUENCE: 11
132   Arg Ser Asp Tyr Lys Phe Tyr Glu Glu Ala Asn Gly Thr Arg Asp His
```

Use of "Xaa" requires mandatory numeric identifiers <220>, <221>, <222>, and <223>. see item 10 on error summary sheet

```
            RAW SEQUENCE LISTING              DATE: 02/15/2001
            PATENT APPLICATION:  US/09/719,379    TIME: 14:29:54

            Input Set : A:\seqlist.txt
            Output Set: N:\CRF3\02152001\I719379.raw

133   1            5              10              15
134   Lys Lys Gly
137 <210> SEQ ID NO: 12
138 <211> LENGTH: 19
139 <212> TYPE: PRT
140 <213> ORGANISM: Haemophilus influenzae strain ntHi-166NP
142 <400> SEQUENCE: 12
143   Arg Ser Asp Tyr Lys Phe Tyr Asn Asp Ala Asn Gly Thr Arg Asp His
144   1            5              10              15
145   Lys Lys Ser
148 <210> SEQ ID NO: 13
149 <211> LENGTH: 19
150 <212> TYPE: PRT
151 <213> ORGANISM: Haemophilus influenzae strain ntHi-1848NP
153 <400> SEQUENCE: 13
154   Arg Ser Asp Tyr Lys Phe Tyr Glu Val Ala Asn Gly Thr Arg Asp His
155   1            5              10              15
156   Lys Lys Ser
159 <210> SEQ ID NO: 14
160 <211> LENGTH: 19
161 <212> TYPE: PRT
162 <213> ORGANISM: Haemophilus influenzae strain ntHi-567
164 <400> SEQUENCE: 14
165   Arg Ser Asp Tyr Lys Phe Tyr Glu Asp Ala Asn Gly Thr Arg Asp Arg
166   1            5              10              15
167   Lys Thr Gly
170 <210> SEQ ID NO: 15
171 <211> LENGTH: 19
172 <212> TYPE: PRT
173 <213> ORGANISM: Haemophilus influenzae strain ntHi-484
175 <400> SEQUENCE: 15
176   Arg Ser Asp Tyr Lys Phe Tyr Glu Asp Ala Asn Gly Thr Arg Lys His
177   1            5              10              15
178   Lys Glu Gly
181 <210> SEQ ID NO: 16
182 <211> LENGTH: 19
183 <212> TYPE: PRT
184 <213> ORGANISM: Haemophilus influenzae strain ntHi-10559RM
186 <400> SEQUENCE: 16
187   Arg Ser Asp Tyr Lys Leu Tyr Glu Val Ala Asn Gly Thr Arg Asp His
188   1            5              10              15
189   Lys Lys Ser
192 <210> SEQ ID NO: 17
193 <211> LENGTH: 19
194 <212> TYPE: PRT
195 <213> ORGANISM: Haemophilus influenzae strain ntHi-601
197 <400> SEQUENCE: 17
198   Arg Ser Asp Tyr Lys Phe Tyr Glu Val Ala Asn Gly Thr Arg Asp His
199   1            5              10              15
```

```
200  Lys Gln Ser
203 <210> SEQ ID NO: 18
204 <211> LENGTH: 19
205 <212> TYPE: PRT
206 <213> ORGANISM: Haemophilus influenzae strain ntHi-226NP
208 <400> SEQUENCE: 18
209  Arg Ser Asp Tyr Lys Phe Tyr Glu Glu Ala Asn Gly Thr Arg Asp His
210  1            5                10              15
211  Lys Arg Ser
214 <210> SEQ ID NO: 19
215 <211> LENGTH: 19
216 <212> TYPE: PRT
217 <213> ORGANISM: Haemophilus influenzae strain ntHi-480
219 <400> SEQUENCE: 19
220  Arg Ser Asp Tyr Lys Phe Tyr Glu Asp Ala Asn Gly Thr Arg Glu Arg
221  1            5                10              15
222  Lys Arg Gly
225 <210> SEQ ID NO: 20
226 <211> LENGTH: 19
227 <212> TYPE: PRT
228 <213> ORGANISM: Haemophilus influenzae strain ntHi-1657MEE
230 <400> SEQUENCE: 20
231  Arg Ser Asp Tyr Lys Phe Tyr Glu Val Ala Asn Gly Thr Arg Glu Arg
232  1            5                10              15
233  Lys Lys Gly
236 <210> SEQ ID NO: 21
237 <211> LENGTH: 19
238 <212> TYPE: PRT
239 <213> ORGANISM: Haemophilus influenzae strain ntHi-214NP
241 <400> SEQUENCE: 21
242  Arg Ser Asp Tyr Lys Phe Tyr Glu Val Pro Asn Gly Thr Arg Asp His
243  1            5                10              15
244  Lys Gln Ser
247 <210> SEQ ID NO: 22
248 <211> LENGTH: 19
249 <212> TYPE: PRT
250 <213> ORGANISM: Haemophilus influenzae strain ntHi-250NP
252 <400> SEQUENCE: 22
253  Arg Ser Asp Tyr Lys Arg Tyr Glu Glu Ala Asn Gly Thr Arg Asn His
254  1            5                10              15
255  Asp Lys Gly
258 <210> SEQ ID NO: 23
259 <211> LENGTH: 19
260 <212> TYPE: PRT
261 <213> ORGANISM: Haemophilus influenzae strain ntHi-499
263 <400> SEQUENCE: 23
264  Arg Ser Asp Tyr Glu Phe Tyr Glu Ala Pro Asn Ser Thr Arg Asp His
265  1            5                10              15
266  Lys Lys Gly
```

```
269 <210> SEQ ID NO: 24
270 <211> LENGTH: 21
271 <212> TYPE: PRT
272 <213> ORGANISM: Haemophilus influenzae strain ntHi-492
274 <400> SEQUENCE: 24
275  Arg Ser Asp Tyr Lys Leu Tyr Asn Lys Asn Ser Ser Asn Ser Thr Leu
276  1               5                   10                  15
277  Lys Asn Leu Gly Glu
278              20
280 <210> SEQ ID NO: 25
281 <211> LENGTH: 21
282 <212> TYPE: PRT
283 <213> ORGANISM: Haemophilus influenzae strain ntHi-502
285 <400> SEQUENCE: 25
286  Arg Ser Asp Tyr Lys Leu Tyr Asp Lys Asn Ser Ser Ser Asn Thr Leu
287  1               5                   10                  15
288  Lys Lys Leu Gly Glu
289              20
291 <210> SEQ ID NO: 26
292 <211> LENGTH: 20
293 <212> TYPE: PRT
294 <213> ORGANISM: Haemophilus influenzae strain ntHi-165NP
296 <400> SEQUENCE: 26
297  Arg Ser Asp Tyr Lys Leu Tyr Asn Lys Asn Ser Ser Asn Thr Leu Lys
298  1               5                   10                  15
299  Asp Leu Gly Glu
300              20
302 <210> SEQ ID NO: 27
303 <211> LENGTH: 20
304 <212> TYPE: PRT
305 <213> ORGANISM: Haemophilus influenzae strain ntHi-495
307 <400> SEQUENCE: 27
308  Arg Ser Asp Tyr Lys Leu Tyr Asn Lys Asn Ser Ser Asp Ala Leu Lys
309   1               5                   10                  15
310  Lys Leu Gly Glu
311              20
313 <210> SEQ ID NO: 28
314 <211> LENGTH: 40
315 <212> TYPE: PRT
316 <213> ORGANISM: LB1 peptide (derived from Haemophilus influenzae)
318 <400> SEQUENCE: 28
319  Arg Ser Asp Tyr Lys Phe Tyr Glu Asp Ala Asn Gly Thr Arg Asp His
320  1               5                   10                  15
321  Lys Lys Gly Pro Ser Leu Lys Leu Leu Ser Leu Ile Lys Gly Val Ile
322              20                  25                  30
323  Val His Arg Leu Glu Gly Val Glu
324          35                  40
326 <210> SEQ ID NO: 29
327 <211> LENGTH: 57
```

<213> Artificial Sequence - see Figure 1

<400> 73
```
ctcttacaca ttccagccct gaaaaagggc atcaaattaa accacacctt aaggaggata      60
taacatatgg atcccatggc cacgtgtgat cagagctcaa ctagtggcca ccatcaccat     120
caccattaat ctagaatcga taagcttcga ccgatgcc                             158
```

<210> 74
<211> 1244
<212> DNA
<213> Haemophilus influenzae - see Figure 2

<220>
<221> CDS
<222> (67)...(1212)

<400> 74
```
ctcttacaca ttccagccct gaaaaagggc atcaaattaa accacacctt aaggaggata      60
taacat atg gat cca aaa act tta gcc ctt tct tta tta gca gct ggc        108
       Met Asp Pro Lys Thr Leu Ala Leu Ser Leu Leu Ala Ala Gly
       1               5                   10

gta cta gca ggt tgt agc agc cat tca tca aat atg gcg aat acc caa       156
Val Leu Ala Gly Cys Ser Ser His Ser Ser Asn Met Ala Asn Thr Gln
 15              20                  25                  30

atg aaa tca gac aaa atc att att gct cac cgt ggt gct agc ggt tat       204
Met Lys Ser Asp Lys Ile Ile Ile Ala His Arg Gly Ala Ser Gly Tyr
             35                  40                  45

tta cca gag cat acg tta gaa tct aaa gca ctt gcg ttt gca caa cag       252
Leu Pro Glu His Thr Leu Glu Ser Lys Ala Leu Ala Phe Ala Gln Gln
         50                  55                  60

gct gat tat tta gag caa gat tta gca atg act aag gat ggt cgt tta       300
Ala Asp Tyr Leu Glu Gln Asp Leu Ala Met Thr Lys Asp Gly Arg Leu
         65                  70                  75

gtg gtt att cac gat cac ttt tta gat ggc ttg act gat gtt gcg aaa       348
Val Val Ile His Asp His Phe Leu Asp Gly Leu Thr Asp Val Ala Lys
     80                  85                  90

aaa ttc cca cat cgt cat cgt aaa gat ggc cgt tac tat gtc atc gac       396
Lys Phe Pro His Arg His Arg Lys Asp Gly Arg Tyr Tyr Val Ile Asp
 95                  100                 105                 110

ttt acc tta aaa gaa att caa agt tta gaa atg aca gaa aac ttt gaa       444
Phe Thr Leu Lys Glu Ile Gln Ser Leu Glu Met Thr Glu Asn Phe Glu
                 115                 120                 125

acc aaa gat ggc aaa caa gcg caa gtt tat cct aat cgt ttc cct ctt       492
Thr Lys Asp Gly Lys Gln Ala Gln Val Tyr Pro Asn Arg Phe Pro Leu
             130                 135                 140

tgg aaa tca cat ttt aga att cat acc ttt gaa gat gaa att gaa ttt       540
Trp Lys Ser His Phe Arg Ile His Thr Phe Glu Asp Glu Ile Glu Phe
         145                 150                 155
```

VERIFICATION SUMMARY   DATE: 02/15/2001
PATENT APPLICATION: US/09/719,379   TIME: 14:29:55

Input Set : A:\seqlist.txt
Output Set: N:\CRF3\02152001\I719379.raw

```
L:14 M:270 C: Current Application Number differs, Replaced Current Application No
L:14 M:271 C: Current Filing Date differs, Replaced Current Filing Date
L:77 M:258 W: Mandatory Feature missing, <220> not found for SEQ ID#:6
L:77 M:258 W: Mandatory Feature missing, <221> not found for SEQ ID#:6
L:77 M:258 W: Mandatory Feature missing, <222> not found for SEQ ID#:6
L:77 M:258 W: Mandatory Feature missing, <223> not found for SEQ ID#:6
L:77 M:340 W: (46) "n" or "Xaa" used: Feature required, for SEQ ID#:6
L:687 M:258 W: Mandatory Feature missing, <220> FEATURE:
L:687 M:258 W: Mandatory Feature missing, <223> OTHER INFORMATION:
```